AD-A081 593 MASSACHUSETTS INST OF TECH LEXINGTON LINCOLN LAB F/G 17/2.1
PREDICTION AND BUFFERING OF DIGITAL SPEECH STREAMS FOR IMPROVED--ETC(U)
NOV 79 W KANTROWITZ, C J WEINSTEIN F19628-80-C-0002
UNCLASSIFIED TN-1979-75 ESD-TR-79-249 NL

1 OF 1
AD
A081593

END
DATE
FILMED
4-80
DDC

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

PREDICTION AND BUFFERING OF DIGITAL SPEECH STREAMS FOR IMPROVED TASI PERFORMANCE ON A DEMAND-ASSIGNED SATELLITE CHANNEL

W. KANTROWITZ
C. J. WEINSTEIN
E. M. HOFSTETTER

Group 24

TECHNICAL NOTE 1979-75

15 NOVEMBER 1979

LEXINGTON MASSACHUSETTS

ABSTRACT

An approach for achieving efficient TASI-like multiplexing of speech on a demand-assigned satellite channel, where a number of ground stations each support a small number of off-hook callers, is proposed and evaluated. The approach presupposes a demand-assignment scheme which allows rapid changes in channel capacity assigned to individual nodes. The components of the approach are: (1) prediction of future speaker activity at each node and channel reservation requests based on the prediction; and (2) buffering of speech at the nodes to aid in prediction and to trade delay for improved multiplexing efficiency.

System performance was evaluated by means of a computer simulation. Results indicate that cutout fraction can potentially be reduced through dynamic allocations based on speaker activity prediction, as compared to the case of fixed allocations. Further improvements can be obtained, at a cost in added delay, by allowing buffering of speech at the individual nodes. These improvements can be taken advantage of by allowing more users to share the channel at a given cutout fraction, or by providing a lower cutout fraction to a fixed number of users.

A

PRECEDING PAGE BLANK-NOT FILMED

CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (CONT'd)

I. INTRODUCTION

Recent trends in integrated voice/data communications network design have begun to favor configurations with a large number of small nodal switches serving small local user groups and with heavy reliance on broadcast satellites for transmission capacity [1,2]. Satellite channel capacity is an expensive commodity in such a system, and flexible Demand Assignment Multiple Access (DAMA) schemes [3] must be relied upon to allocate this commodity efficiently according to fluctuating demands from the various earth terminals. Channel utilization can be significantly enhanced through the application of TASI (Time Assigned Speech Interpolation) [4] or DSI (Digital Speech Interpolation) [5] wherein off-hook voice callers occupy channel capacity only during talkspurts and not during silence periods. Since talkers in conversation are typically silent more than 50% of the time [6,7], the potential capacity saving, generally referred to as the "TASI advantage" is greater than a factor of two. However, with standard TASI or DSI systems, achievement of the full potential TASI advantage requires that a large number of talkers (typically 50 or more) be statistically multiplexed at a particular node. The configuration of concern here is a number of small earth stations or nodes, where the number of off-hook callers at each node is too small to achieve efficient TASI multiplexing by standard techniques, but where the aggregate number of callers sharing the satellite would be large enough for efficient multiplexing if all the users were located at one node. This paper describes and evaluates a proposed approach for achieving

efficient TASI-like multiplexing in this configuration. The approach presupposes a DAMA scheme such as Priority-Oriented Demand Assignment (PODA) [3] which allows stations to request and rapidly obtain changes in their share of a Time-Division Multiple Access (TDMA) channel. The components of the approach are:

1. prediction [8] of the number of callers in talkspurt at each station ahead by the time (minimum of one satellite round-trip propagation delay) required to change channel capacity allocation, combined with requests for channel capacity on the basis of this prediction;
2. variable-length buffering of speech at each station and trading of delay for TASI advantage [9].

The results, which are obtained primarily through computer simulation, show that this dual approach provides substantial potential improvement in TASI advantage over a system with channel allocations which cannot be changed rapidly enough to respond to talkspurt/silence variations and without variable buffering at the nodes. Note that the term "TASI advantage" is used generically in this paper to refer to the ratio of the number of off-hook callers to the system channel capacity, where one unit of capacity is taken to be just sufficient to support one caller during talkspurt. The use of this term should not confuse the fact that the system under consideration is quite different from the classical TASI system.

The organization of the paper is as follows. Section II details the system model and strategies for improved TASI performance and discusses

the interplay among: talkspurt activity predict-ahead time, DAMA system response time, speech delay, and margin to be associated with channel capacity requests. Section III describes the theoretical basis of speaker activity prediction and presents an example illustrating the potential benefits of prediction. Section IV deals with the TASI performance improvement results that can be achieved with speaker activity prediction but without flexible buffering at the nodes. Section V discusses the combined effects of prediction and buffering. Section VI deals with the sensitivity of the simulation results to various system parameters, and Section VII summarizes potential TASI performance improvements for the strategies considered. Conclusions are summarized in Section VIII.

II. SYSTEM MODEL AND STRATEGIES FOR IMPROVED TASI PERFORMANCE

The multi-node satellite-based communication system model of interest here is depicted in Fig. 1. There are N ground stations, and the nodal processor at the ith station supports M_i off-hook callers. Functions of the nodal processors include multiplexing and demultiplexing of local traffic, as well as the processing necessary to support the satellite demand assignment algorithm. Application of Speech Activity Detection (SAD) and transmission only during talkspurt is assumed for each caller so that the transmission rate which must be supported at a node varies with the number of active talkspurts.

The satellite channel capacity is assumed to be shared among the N nodes on a dynamically demand-assigned burst-TDMA basis. The capacity allocated to each individual station is assumed to be in the form of a

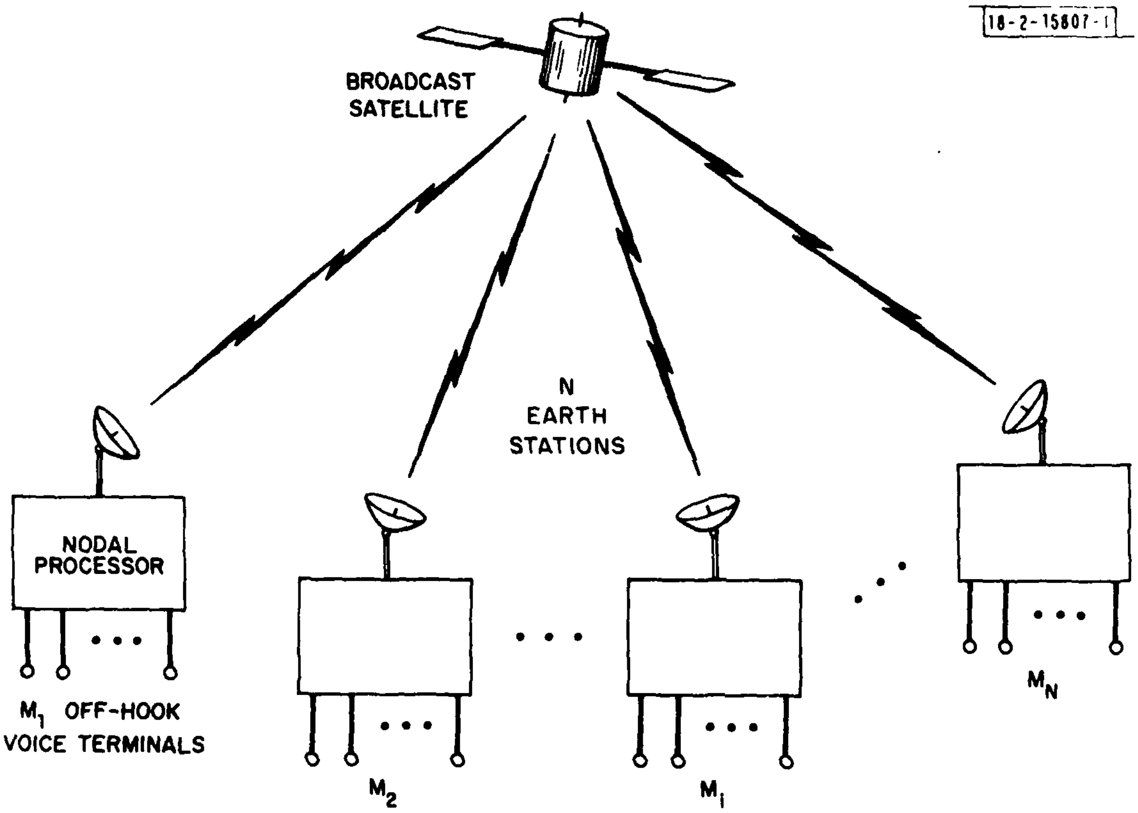

Fig.1. Configuration of multi-node satellite communications system.

variable-size "stream" [3]. Once every T_S sec the station has the opportunity to transmit a burst segment, where the maximum number of bits in this segment is the stream size. The DAMA algorithm is assumed to schedule these burst segments to be transmitted from the individual stations in a non-interfering and efficient manner. To minimize end-to-end delays it is desirable that T_S be kept as short as possible, on the order of 20-40 ms. It is not necessary that T_S match the frame interval which is associated with the TDMA pattern of the DAMA algorithm. Each segment, as shown in Fig. 2, is assumed to contain a short reservation request slot used to request changes in the size of the stream plus a set of speech slots each capable of carrying the amount of digital speech produced by one active voice terminal during one frame interval. For simplicity of simulation and analysis it is assumed here that all speech slots are of equal size, although the strategies and general nature of the results are not limited to this case. The nodal DAMA processor inserts into each reservation slot a request for a number of speech slots. This request number may vary slowly on the basis of variations in the number of off-hook callers M_i or more rapidly on the basis of variations in the number of callers in talkspurt. In either case the request cannot be granted until it has been received by all stations, at least one satellite round-trip time (≈270 msec) after it is issued. A distributed DAMA algorithm is assumed wherein the nodal processors at each station collect all requests and allocate speech slots based on identical, fair round-robin algorithms [10]. Generally this channel allocation might occur synchronously with the frame structure

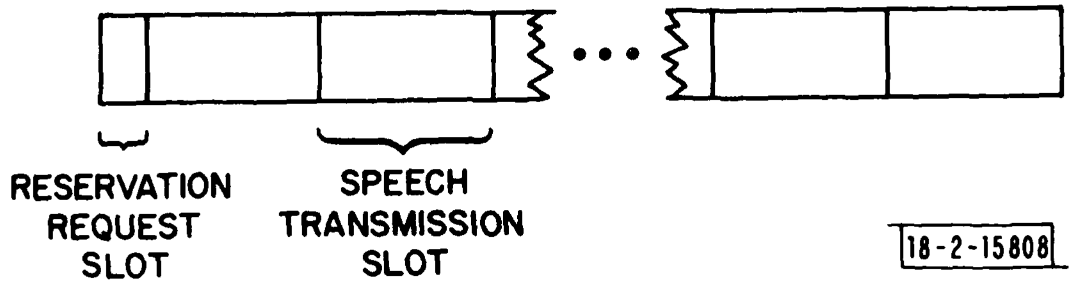

Fig.2. Format of burst segments transmitted in a single multiplexed speech stream. The reservation of stream capacity allows a node to transmit one of these segments every T_S sec. The stream size (number of speech slots) may be varied dynamically by changes in the reservation request.

of the DAMA algorithm. For convenience in most of the simulation work here it has been assumed that allocations of channel capacity are updated every T_S sec, upon receipt of new reservation requests from all stations. Since T_S is typically much shorter than a satellite round-trip time, a "reservation pipeline" is formed wherein a number of reservation requests are propagating across the transmission link at any time.

The reservation response delay is not a significant limiting factor in responding to call initiations or terminations, where response times on the order of seconds are acceptable. However, achievement of efficient TASI multiplexing without nodal buffering in the case of a small number of callers per node requires that each node's slot allocation closely match the number of active (i.e., currently in talkspurt) speakers at that node. Because of the reservation response delay, the best each node can do to achieve this match is to issue slot requests based on a prediction of the number of talkers likely to be active one reservation response delay in the future. If, due to inaccurate prediction or limited overall satellite capacity, a node's slot allocation at a particular time becomes temporarily insufficient to support the instantaneous number of active talkspurts, then the overflow speech must either be discarded immediately, or buffered (adding delay) at the node until transmission capacity becomes available or the buffer overflows. Both cases are considered here.

The strategies considered here can apply whether a packet- [11,12] or circuit-oriented [5] transmission format is utilized for the digital speech. As discussed in [13], the required control overhead for packet

transmission can be reduced to a level comparable with that required to accommodate talker activity information in a circuit-switched system, if fixed virtual-circuit routing [14] is used for the packets. The primary remaining difference then becomes the flexible buffering allowed by the asynchronous nature of the packet system. However a digital circuit-switched DSI system can also be augmented to include flexible buffering [15]. For convenience, the term "packet" will be used here to denote the speech information which is accommodated in a speech slot (see Fig. 2), and speech buffers (when applied) are assumed to accommodate packet-sized units. However, it should be understood that the strategies and results are not strictly limited to packet systems.

A block diagram of the functions to be carried out at each node is shown in Fig. 3. The off-hook voice terminals transmit digital speech packets (during talkspurts only) through the multiplexer which feeds a multiplexed speech stream into the buffer. Once every stream interval T_S, the speech stream transmitter discharges from the buffer the number of packets that can fit in its current stream segment. The maximum time T_B that a packet is allowed to remain in the buffer is set by a delay control parameter. Packets not discharged within this time are discarded. The cutout fraction, defined as percentage of packets discarded, is a key performance parameter in the system. Generally cutout fractions less than 0.5% will be essentially unnoticeable to users and cutout fractions on the order of 1% can be tolerated without significant degradation in user acceptability. This holds both for standard TASI systems [4] where

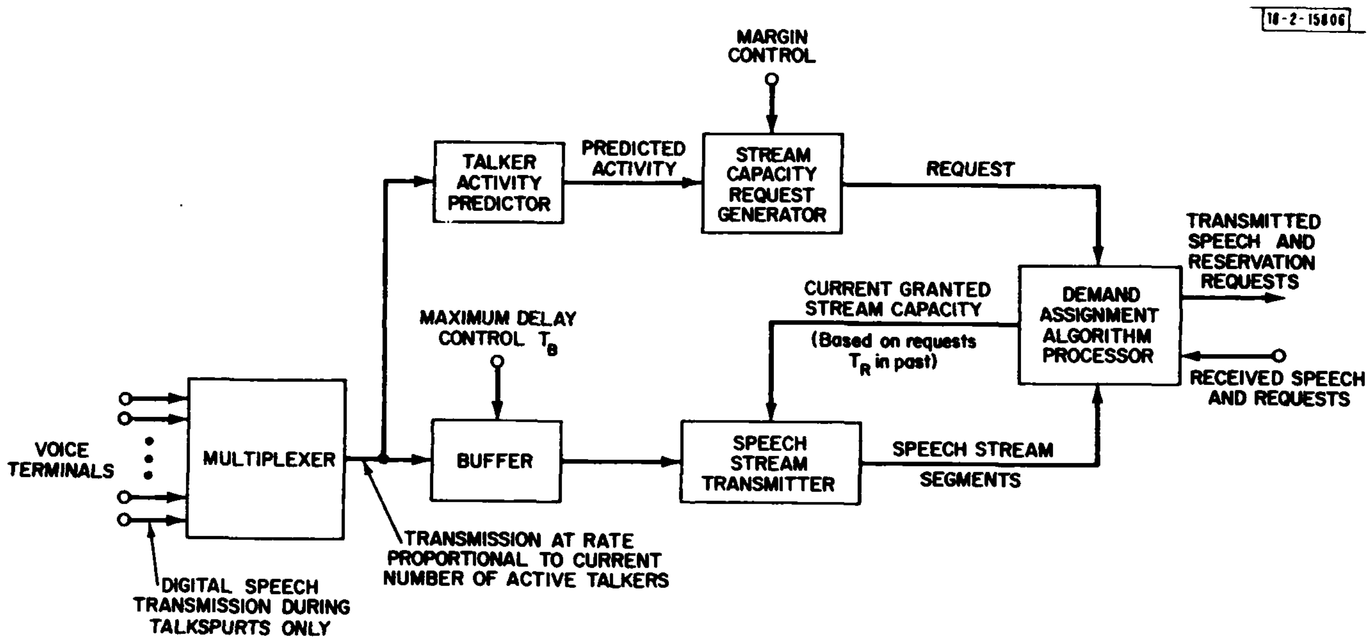

Fig.3. Block diagram of functions to be carried out at each node. Functions include multiplexing, buffering, transmission, prediction, reservation request generation, and execution of distributed demand assignment algorithm.

cutouts occur only at talkspurt onsets, and for the system under consideration here where speech loss can be dispersed [16] through any part of a talkspurt. Minimal buffering delay (corresponding to a standard synchronous TASI or DSI system) results when the delay control parameter is set such that no packet remains in the buffer longer than one stream interval. Stream capacity is granted by the DAMA algorichm on the basis of the reservation requests most recently received from all stations and processed by the DAMA algorithm. The speech activity predictor observes the current number of active talkers in the multiplexed speech stream at the buffer input, and estimates the number of talkers likely to be active at a predict-ahead time T_P into the future. The request algorithm adds a margin M_A to this prediction to produce a reservation request for transmission along with the current speech frame. Margin is chosen (as discussed in more detail below) in order to balance optimally for a given overall satellite load, packet losses due to (1) insufficient reservation request by the individual node; and (2) denial of reservation requests by the DAMA algorithm when the sum of all nodal requests exceeds channel capacity.

There are fundamental interrelationships in this system among the maximum buffer delay T_B, the required predict-ahead time T_P, the reservation response time T_R, and the margin M_A. The growing uncertainty of predicting further into the future implies that M_A should increase with T_P. If T_B is set to zero, then T_P must equal T_R, which is lower-bounded by the satellite round-trip time. On the other hand an increase in T_B has

the effect of producing a corresponding decrease in T_P. In particular, if $T_B = T_R$ then no prediction is necessary because the speech can be buffered locally just long enough to make the desired change in channel allocation. The TASI performance of the overall satellite system for this special case will be as effective as if all callers were multiplexed at a single node. The cost for obtaining this multiplexing performance is an added delay of T_R. The potential benefit of speaker activity prediction is to reduce this delay while still achieving efficient channel utilization.

III. SPEAKER ACTIVITY PREDICTION

Consider M independent off-hook callers each alternating between talkspurt (active mode) and silence (inactive) and let $n(x)$ denote the number of active talkers at time x. The optimum, least squares prediction of $n(t + \tau)$, given the past history of $n(x)$ through time t, is the conditional expectation $E[n(t + \tau)|n(x), -\infty < x \leq t]$. Assume a model of talkspurt and silence durations as exponentially-distributed random variables with means μ^{-1} and λ^{-1}, respectively. This implies that $n(x)$ is a Markov process [17] so that this conditional expectation is the same as $E[n(t + \tau)|n(t)]$; i.e., all knowledge of the past of $n(x)$ prior to t is summarized in the current value $n(t)$.

With the Markov assumption, an explicit expression for the conditional probability distribution of $n(t + \tau)$ given $n(t)$ can be obtained [8,18], from which the optimum predictor (the conditional mean) and its variance are easily derived. The talkspurt/silence behavior of each talker can be represented as a two-state Markov process where the state

transition rate from silence to talkspurt is λ and the corresponding talkspurt to silence transition rate is μ. Let p_1 represents the probability that a particular talker is active at $t + \tau$ given that talker was active at t, and let p_2 represent the probability that a talker is active at $t + \tau$ given that talker was inactive at t. Then standard Markov process analysis techniques [19] applied to this two-state model lead to the following results:

$$p_1 = (1 + \frac{\mu}{\lambda} e^{-(\lambda+\mu)\tau})/(1 + \frac{\mu}{\lambda}) \tag{1}$$

$$p_2 = (1 - e^{-(\lambda+\mu)\tau})/(1 + \frac{\mu}{\lambda}). \tag{2}$$

If $j = n(t)$ then the probability that x of these j previously active talkers will still be active at $t + \tau$ is then given by the binomial function

$$b(x, j, p_1) = \binom{j}{x} p_1^{x}(1 - p_1)^{j-x}. \tag{3}$$

Similarly the probability that y of the M-j previously inactive talkers will be active at $t + \tau$ is

$$b(y, M-j, p_2) = \binom{M-j}{y} p_2^{y}(1 - p_2)^{M-j-y}. \tag{4}$$

The total number of active talkers at $t + \tau$ is just $k = n(t+\tau) = x + y$ and the desired conditional distribution for k is simply the convolution of the two independent binomial distributions (3) and (4). The optimum predictor is the sum of the means of the constituent binomial distribution,

$$\mu_j(\tau) \equiv E[n(t+\tau)|n(t)=j] = jp_1 + (M-j)p_2$$

$$= \frac{M}{1+\frac{\mu}{\lambda}} \left\{1 - [1-(1+\frac{\mu}{\lambda})\frac{j}{M}]e^{-(\lambda+\mu)\tau}\right\} \quad . \qquad (5)$$

Similarly the mean-squared error of the optimum predictor is the sum of the variances of the independent constituent binomial distributions,

$$\sigma_j^2(\tau) \equiv E\left\{[n(t+\tau) - \mu_j(\tau)]^2|n(t) = j\right\}$$

$$= j\ p_1(1-p_1) + (M-j)p_2(1-p_2)$$

$$= \frac{M\frac{\mu}{\lambda}}{(1+\frac{\mu}{\lambda})^2} [1-e^{-(\lambda+\mu)\tau}] \left\{1 + [\frac{j}{M}(\frac{\mu}{\lambda} - \frac{\lambda}{\mu}) + \frac{\lambda}{\mu}]\ e^{-(\lambda+\mu)\tau}\right\} \quad (6)$$

Plots of (5) and (6) for the case $M = 10$, $\mu^{-1} = \lambda^{-1} = 1.5$ sec are shown in Figs. 4 and 5. Note that when $\lambda = \mu$, $\sigma_j^2(\tau)$ is independent of j. Inspection of these curves indicates that reasonably good prediction (±1 speaker rms error) can be realized for prediction times on the order of a round-trip satellite delay. Thus the predictability of the speaker activity process, which results from time correlation due to typical

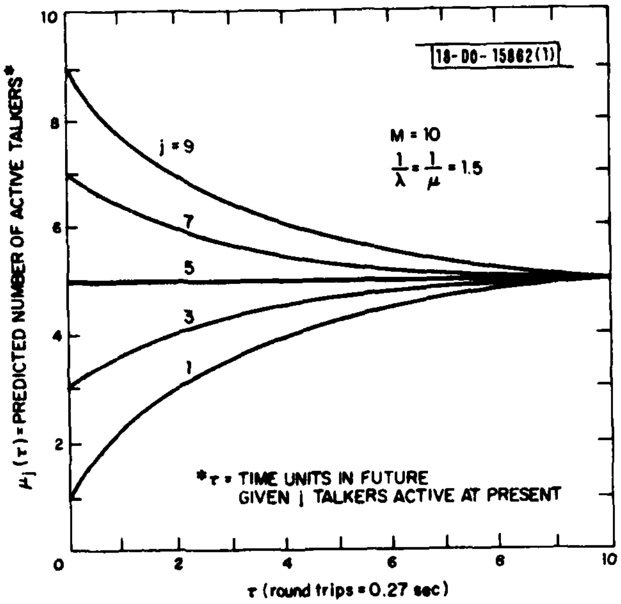

Fig.4. Optimum talker activity predictor.

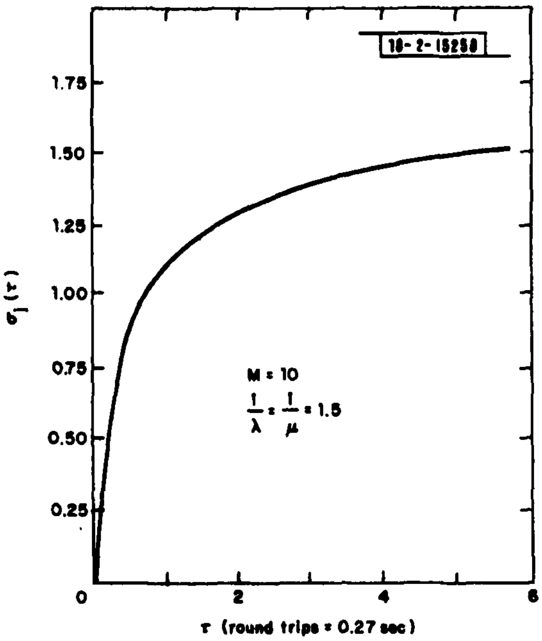

Fig.5. Prediction error of optimum talker activity predictor.

talkspurt and silence durations, seems to offer potential for TASI advantage improvements along the lines discussed above.

The sensitivity of the above results to the Markov assumptions made above was tested by comparing the theoretical results to results obtained using a computer program simulating the behavior of a number of independent talkers following measured [6], non-exponential talkspurt/ silence duration distributions. For M close to 10, measurements of the conditional mean and variance of future speaker activity closely matched the theoretical results given in (5) and (6).

A graphical illustration of the potential benefits of speaker activity prediction is shown in Fig. 6. The identical solid curves in the top and bottom parts of the figure represent an 8-second segment of a talker activity time function $n(t)$ obtained from simulation with $M = 10$ and exponential talkspurt/silence distributions. The average talkspurt duration was 1.23 sec, the average silence duration was 1.34 sec, the corresponding fractional talker activity $p = .48$, and the average talker activity $\overline{n(t)} = Mp = 4.8$. The dotted curves represent channel allocation $c(t)$ in slots/frame. The bottom part of the figure corresponds to a fixed allocation $c(t) = 6$. Dark gray areas indicate periods where $n(t) < 6$ so that capacity is wasted. Light gray areas indicate periods where $n(t) > 6$ and where, assuming no buffering, speech packets will be discarded. In the top curve $c(t)$ was obtained by predicting $n(t)$ 280 msec into the future and adding sufficient margin so that the average $\overline{c(t)}$ equals 6. It is apparent that the predictor, while far from perfect, does tend to track

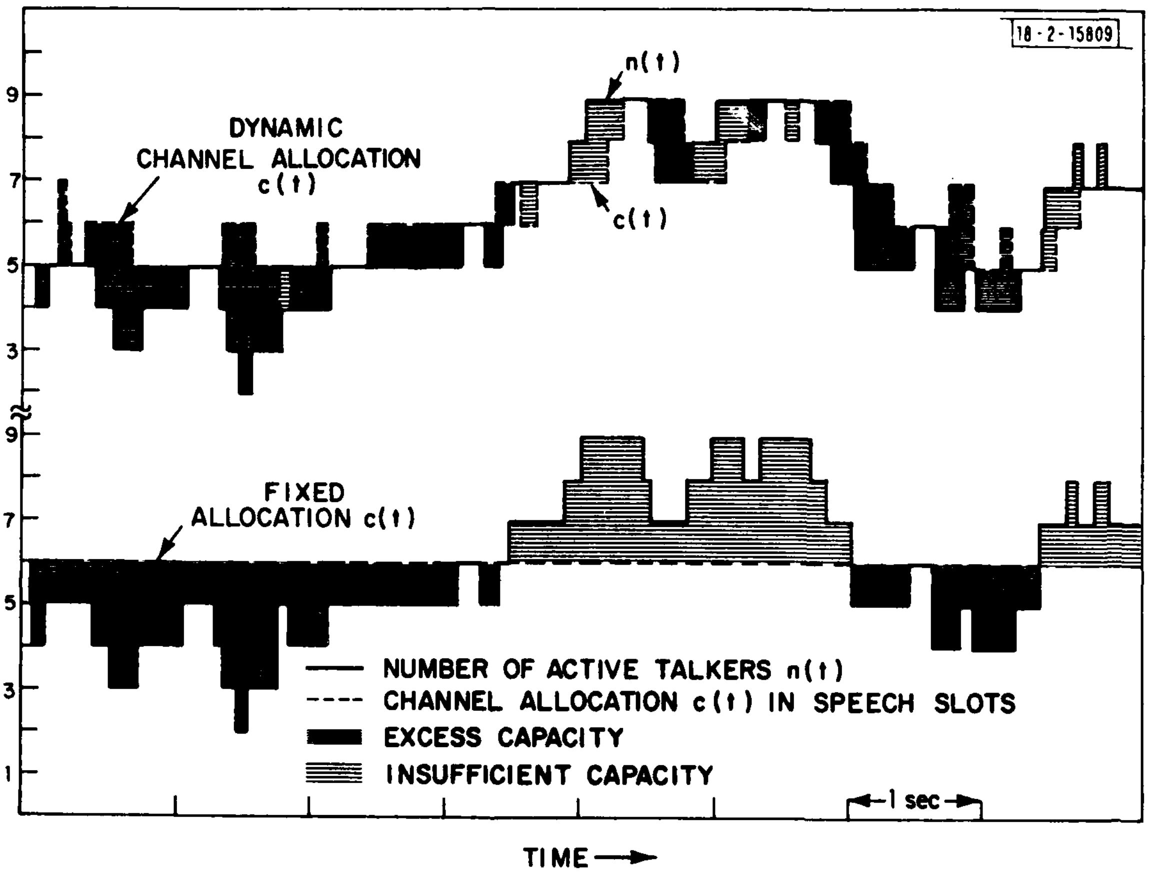

Fig.6. Sample functions of active talker process n(t) and channel allocation c(t) illustrating comparison of fixed allocation with prediction-driven dynamic channel allocation.

the changing talker activity. Both packet loss and wasted capacity are substantially reduced for this example with predictor-based allocations as compared to a fixed allocation with the same long-term average.

In the next section simulation results for TASI performance with prediction-driven stream reservations but without buffering at the nodes will be described. The complexity of the system made simulation the most viable approach, particularly when buffering was combined with prediction. However some analysis is possible, and the development of a complete analytical model for the system is an interesting subject for further research. As an example of an analytic approach that can be followed, consider the calculation of packet loss for M talkers at a node (following the Markov talkspurt/silence model) with prediction-based reservation requests, where interaction with other nodes is ignored such that all requests are assumed to be granted. With j talkers active at t, the channel allocation at $t + \tau$ depends only on j. The typical channel allocation function, including margin, which was used in this study is

$$c_j = \min\{I[\mu_j(\tau) + M_A], M\} \tag{7}$$

where I denotes integer part. The interpretation of (7) is that the reservation request is obtained by adding margin and taking the integer part so that an integral number of slots are requested; in addition the request never is allowed to exceed the number of off-hook callers M. The cutout fraction is

$$\phi = \frac{1}{Mp} \sum_{j=0}^{M} p_j(j) \sum_{k=c_j+1}^{M} [k - c_j] p_{k|j}(k|j) \tag{8}$$

where $p_j(j) = b(j,M,p)$ is the unconditional probability of j active talkers and the conditional distribution $p_{k|j}(k|j)$ is obtained from a convolution of two binomials as discussed above. This calculation is straightforward, but interaction with other ground stations and the introduction of buffering make the problem of analytically determining cutout fraction for the multi-node system significantly more complex.

IV. TASI PERFORMANCE IMPROVEMENTS WITH PREDICTION-DRIVEN STREAM RESERVATIONS

In this section, simulation results on system performance with prediction but without additional buffering delay at the individual nodes are presented. Referring to Fig. 3, the constraint applied is that speech packets which are not transmitted within the inter-packet interval T_S are discarded. The primary performance measure of the system is cutout fraction. A key issue was the selection of the correct margin level to minimize this loss fraction.

An illustration of the nature of the simulation results as well as a discussion of the key system variables can be carried out in the context of the example shown in Fig. 7. Here the variation of packet loss with margin is presented for the case of N = 12 nodes, M = 10 off-hook talkers per node, and an overall satellite capacity assumed to be sufficient to accommodate 80 voices in talkspurt. The system TASI advantage, or ratio

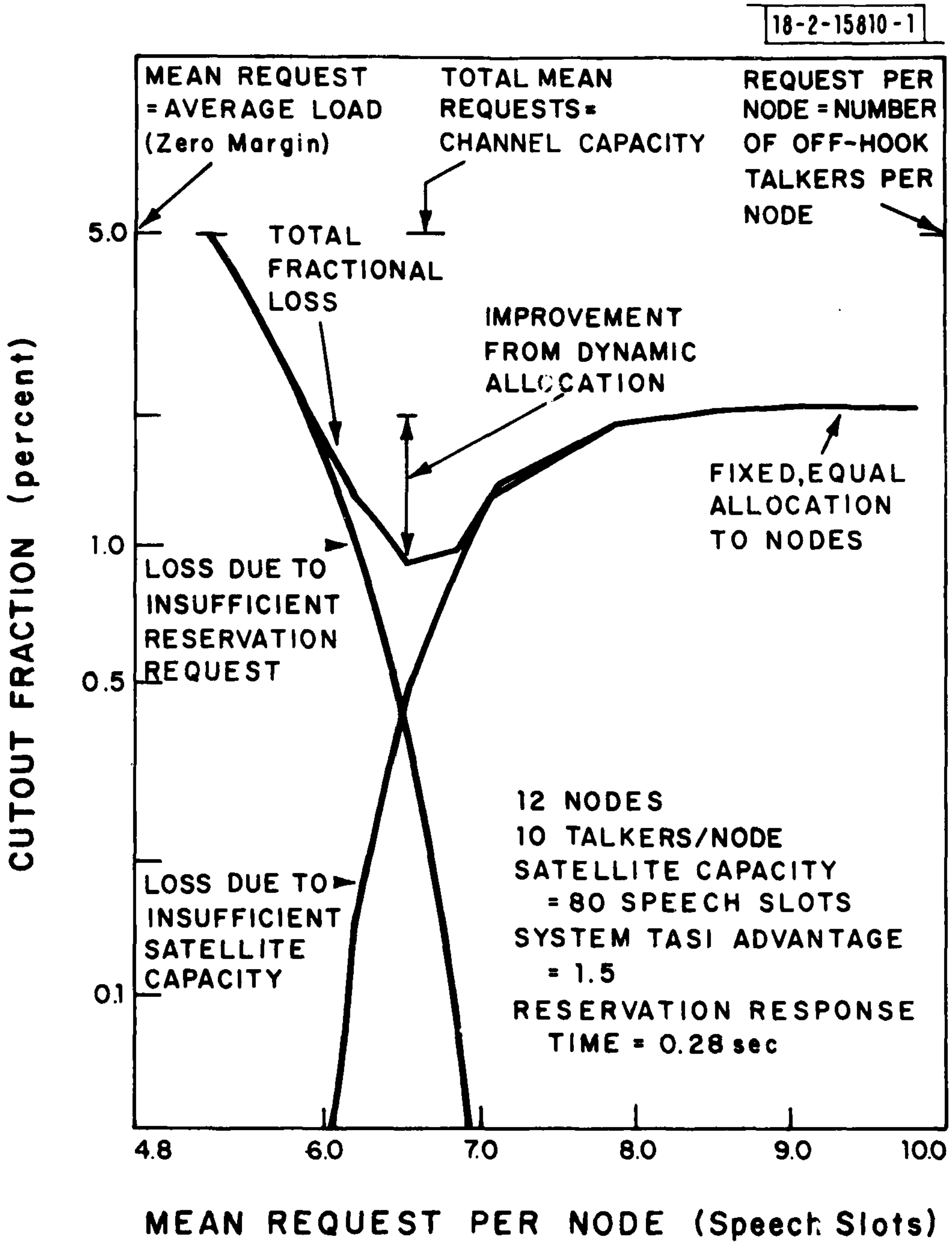

Fig.7. Example of the effects of margin and of potential performance improvement with dynamic allocations.

of number of off-hook callers to channel capacity, is 1.5. In this, as in most of the runs, callers were assumed to generate one packet every T_S = 20 ms during talkspurt. During the runs, all pertinent system variables and statistics are generally updated every T_S sec. Talkspurt and silence durations were generated randomly from exponential distributions with means of 1.23 and 1.34 sec, respectively, for a talker activity fraction of 0.48. Each plotted point represents an average cutout fraction over 1200 sec of simulated real-time activity; this duration was found to be more than sufficient to obtain statistically stable results. Each station updates its prediction on the basis of the current number of local active talkers and issues a new reservation request every T_S sec. The system reservation response time T_R, which for this case is equal to the required predict-ahead time T_P, is taken as 0.28 sec, just slightly longer than the satellite round-trip time. The manner is which reservation requests are generated from prediction and margin is indicated in equation (7). In order to provide more insight into system behavior the choice has been made to plot fractional loss as a function of mean reservation request per node rather than directly as a function of margin M_A. Clearly mean request increases with margin to a maximum of 10 slots/node.

As shown in the figure, cutout fraction can be divided into components arising from two causes: insufficient reservation request at the individual station and denied reservation requests because satellite capacity was insufficient to accommodate all requests. For low mean request (and margin) almost all the loss is due to the first cause.

As mean request per node increases, loss due to insufficient satellite capacity becomes dominant. Overall cutout fraction is minimized when these effects are balanced in an optimal way. For this example the optimal mean request is about 6.6 slots/node. For the 12-node system, with an overall capacity of 80 slots, optimal performance is achieved when the overall average requested capacity is approximately equal to the total channel capacity. A mean request of 10 slots/node corresponds to the case where prediction is essentially ignored and each station always requests enough capacity to accommodate all 10 talkers. In this case the round-robin DAMA algorithm will provide equal allocations to all nodes. The 2.0% packet loss for the case of equal allocations should be compared to the minimum loss of 0.9%. This graphically shows the potential improvement due to prediction-driven dynamic allocation with the correct choice of margin as compared to equal allocation. An assumption which has been made in this work is that nodes are granted capacity only up to the amount they request, even if not all slots on the satellite channel are requested at a particular time. This excess capacity could be utilized by other traffic (e.g., data) on the channel. If no other traffic is present, then even for the case of optimal margin a small percentage of the available slots on the satellite channel is wasted because no node requests them. It has recently been shown [20] that a small degree of further performance improvement can be achieved by distributing unrequested slots among the nodes.

Performance curves similar to Fig. 7 were obtained for numbers of nodes varying from 8 to 16, with all other system parameters kept the

same. Fig. 8 plots the mean reservation request per node minimizing cutout fraction as a function of the number of nodes. The observation that the margin should be chosen such that the total mean reservation request is roughly equal to the channel capacity is shown to hold for all cases. Fig. 9 compares percentage loss with variable allocation as determined by prediction with optimal margin against percentage loss with fixed allocations. The improved performance over the range of system TASI advantage is as illustrated.

The required predict-ahead interval (assumed to be .28 sec in Figs. 7-9) is a key parameter of this system. The further into the future one must predict, the less accurate prediction becomes and the less advantage can be obtained. Fig. 10 displays a family of curves, each for a different predict-ahead interval, showing the percentage packet loss at various TASI advantages. The case of equal allocations is included for reference; this can be considered as corresponding to an infinite predict-ahead interval since no improvement from prediction is possible. It should be noted that unless buffering delay is allowed at the nodes, the actual required predict-ahead interval must exceed the satellite round-trip time of 0.27 sec.

The optimal margins in the results shown so far were determined by carrying out a number of runs with different but fixed values of margin and empirically determining an optimum. An investigation was carried out to determine if margin could be adapted automatically to system conditions (number of nodes, number of talkers, etc.). To zero in on optimal margin

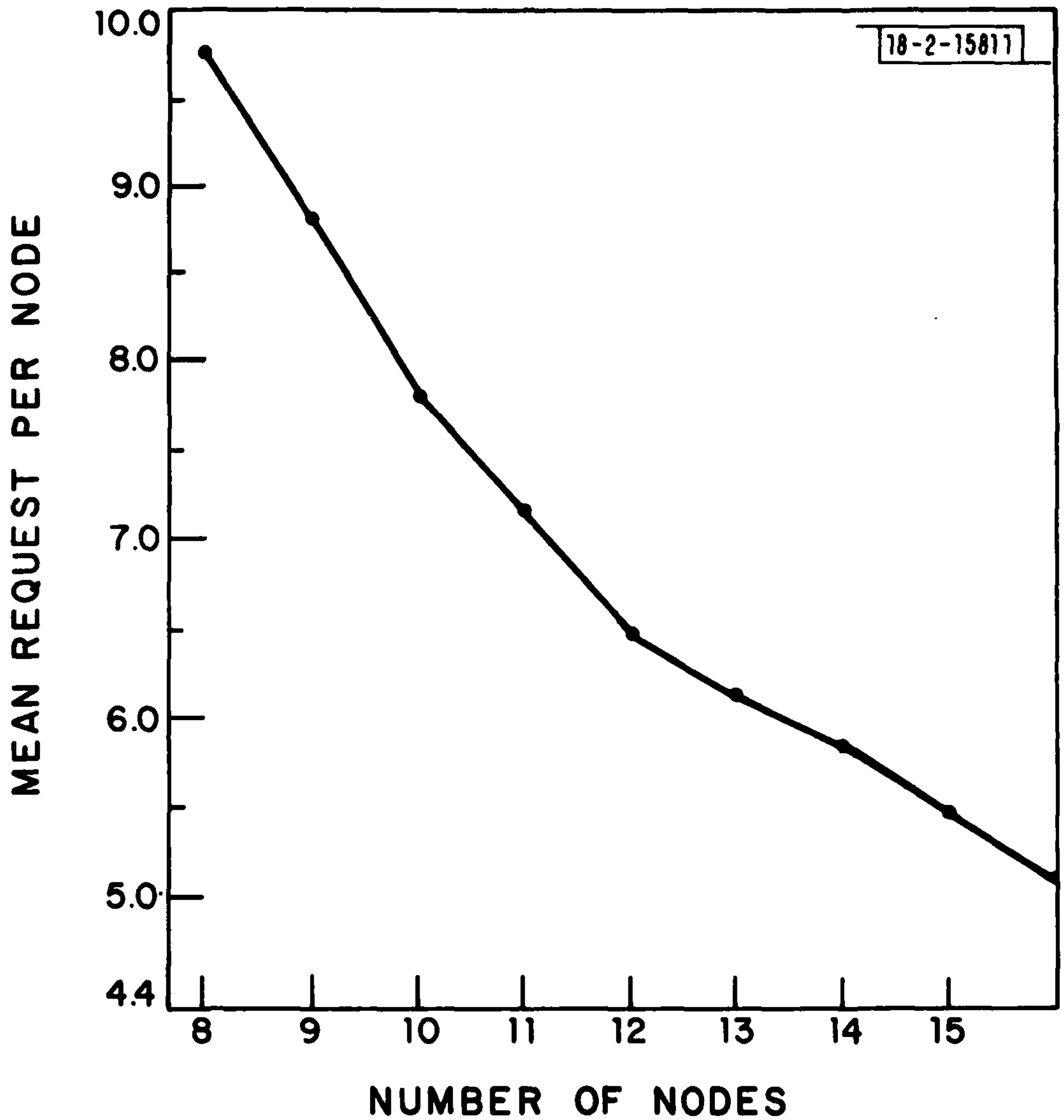

Fig.8. Mean request per node yielding smallest fractional speech loss, as a function of number of nodes. Results indicate that margin should be chosen such that total mean reservation request is approximately equal to satellite channel capacity.

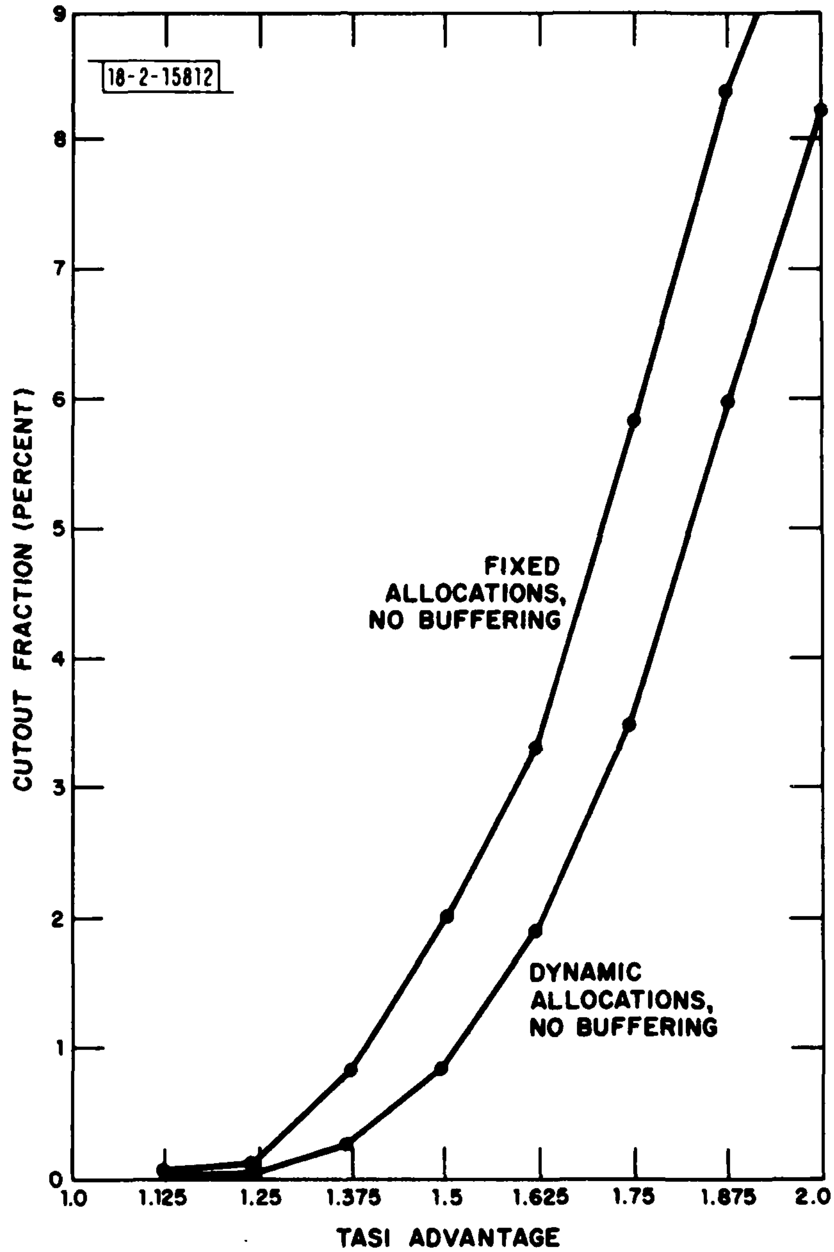

Fig.9. Comparison of cutout fraction with fixed and dynamic allocations, as a function of system TASI advantage. Referring to Fig.8, note that system TASI advantage varies from 1.0 to 2.0 as the number of nodes varies from 8 to 16.

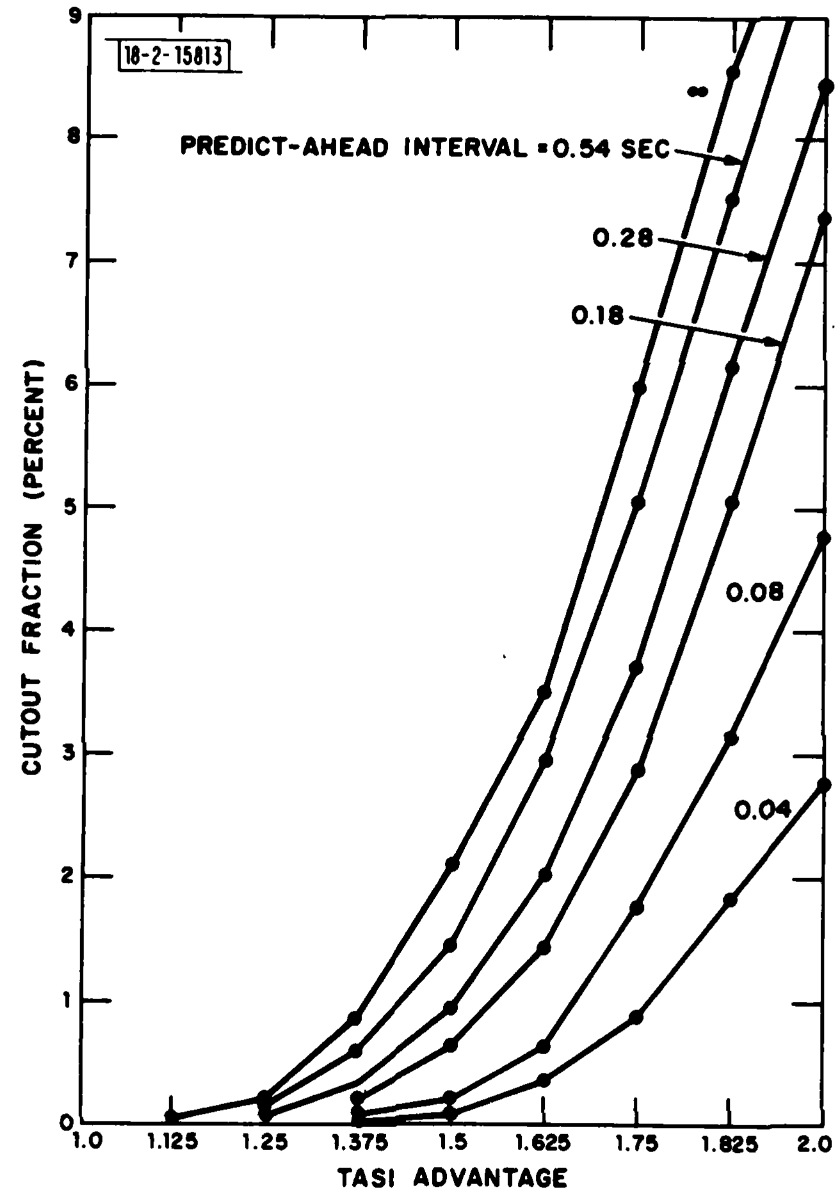

Fig.10. Cutout fraction as a function of system TASI advantage for various predict-ahead intervals.

the result was applied that the total mean reservation request should be close to channel capacity. Each node was allowed to observe the total number of reservation requests currently being made, and then to make incremental adjustments in its own margin (and mean request level) to bring the total request level closer to channel capacity. The results were quite encouraging. The nodes quickly approached the optimal margin and stayed at or near this value with small oscillation. Percentage packet loss was very close to the results obtained with optimal fixed margins.

V. TASI PERFORMANCE IMPROVEMENTS WITH COMBINED PREDICTION AND SPEECH STREAM BUFFERING

As shown in the previous section, dynamic allocations based on prediction can improve system performance, decreasing percentage packet loss for a given TASI advantage. Further improvements are possible if buffering is allowed at the nodes. Buffering can avoid packet loss during temporary overload conditions and can effectively reduce required predict-ahead interval. The advantages of buffering for the case of a single multiplexer with fixed channel capacity are discussed in [9].

For the multinode system considered here, the effects of both fixed delay and variable delay buffering have been examined. In fixed delay buffering, each speech packet is held in a buffer at the transmitting node for a fixed period of time. When this time has elapsed the packet is transmitted if there is sufficient allocation or discarded otherwise. Fixed delay results in a direct reduction of required predict-ahead

interval by the length of the delay. As shown in Fig. 10, smaller predict-ahead intervals result in more accurate prediction and lower percentage packet loss. As an example, refer to Fig. 10 and consider a TASI advantage of 1.625. When prediction .28 sec into the future is required to match the system reservation response time, there is a 2% packet loss. However, a 0.2 sec fixed delay reduces predict-ahead time to .08 sec for the same reservation response time, and reduces packet loss to 0.61%. Of course, the users must tolerate the increase in speech delay.

For the case of variable delay, the buffer is also limited to a fixed maximum delay but packets stay in the buffer only as long as necessary. Buffer size and delay tend to grow when many talkers are active and diminish when many talkers are silent. Variable delay also tends to decrease the required predict-ahead interval but the relationship is not as direct as with fixed delay. However, the need for optimal prediction and margin is not as crucial for the case of variable delay since the buffer tends to smooth out momentary mismatches.

Figure 11 summarizes simulations that have been run to measure the inter-relationship and performance improvement gained from combinations of fixed and variable allocation in conjunction with fixed-delay and variable-delay buffering. System parameters not given explicitly are as in Figs. 7-9. For comparison purposes the results with no buffering delay and fixed allocation are shown. Buffer limits of 100 ms and 200 ms were considered. For each buffer size, progressively improving performance resulted for the following three cases: (1) fixed channel allocation,

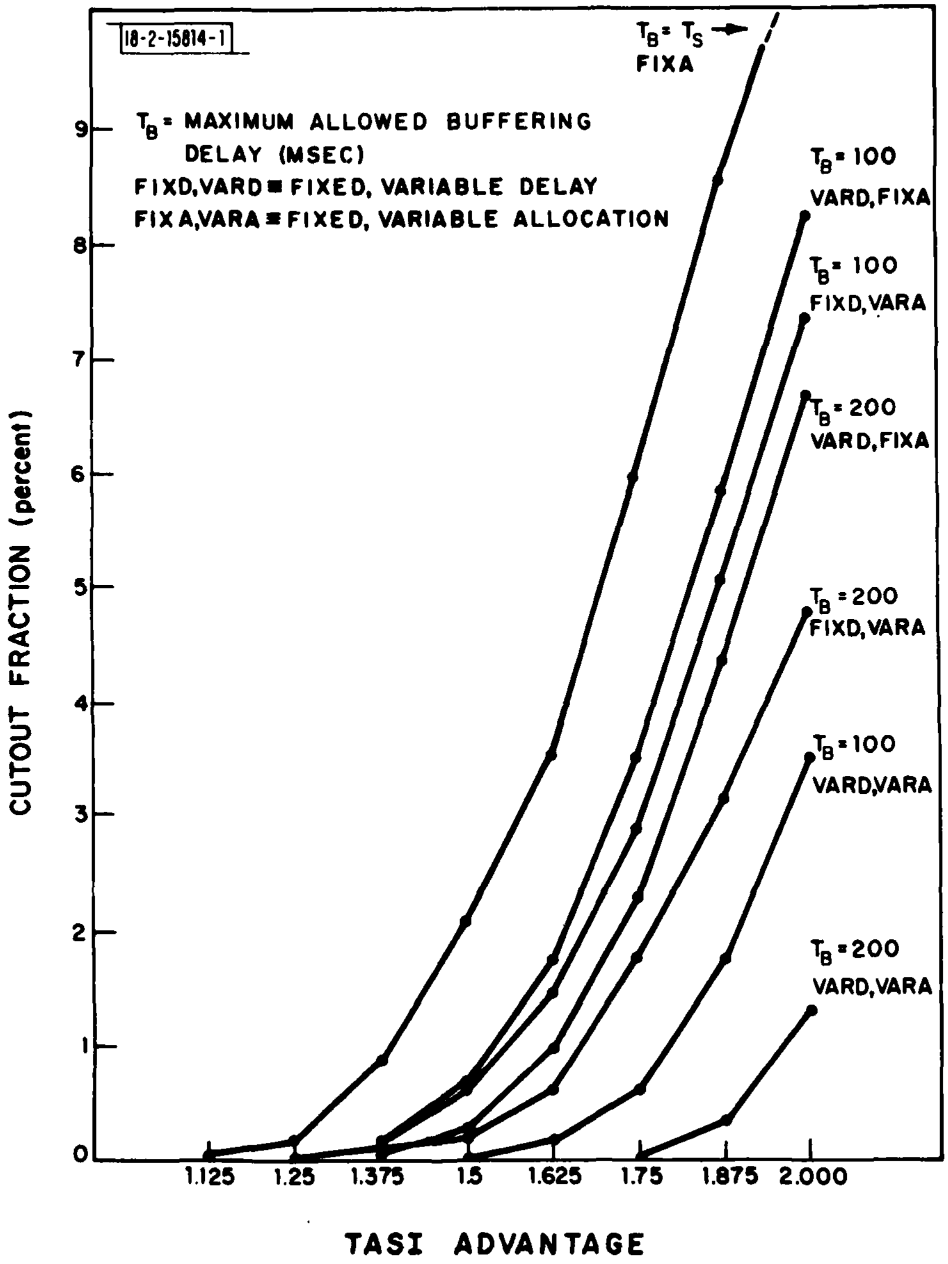

Fig.11. Cutout fraction as a function of TASI advantage for various combinations of buffering and allocation stragegies.

variable delay; (2) fixed delay, variable channel allocation; (3) variable delay, variable channel allocation. For the case of a 200 msec variable delay with variable allocation, the packet loss performance is excellent, in that the system could be run at a TASI advantage of approximately 1.9 with only 0.5% packet loss.

VI. SENSITIVITY ANALYSIS

In most of the simulation runs described above, a large number of the system variables were kept fixed. A number of additional runs have been carried out to explore the sensitivity of the results to these parameters.

The speaker activity model was central to this work and an investigation of sensitivity to its details were explored. The exponential talkspurt and silence duration distributions were replaced by distribution drawn from Brady's [6] data. As long as the mean talkspurt and silence duration were maintained the same, the differences in results were negligible. The effects of different values of mean talkspurt and silence durations was investigated, and an example of the results is shown in Fig. 12. Doubling the durations while keeping activity fraction constant is seen to improve performance. This is due to increased correlation and predictability of the talker activity process. A lower activity fraction is also seen to result in improved performance. It is likely that typical speech communication on satellite channels [7] is characterized by lower activity fraction and longer talkspurt and pause durations than used in obtaining most of the results reported here. Therefore the results presented here should be somewhat conservative in this respect.

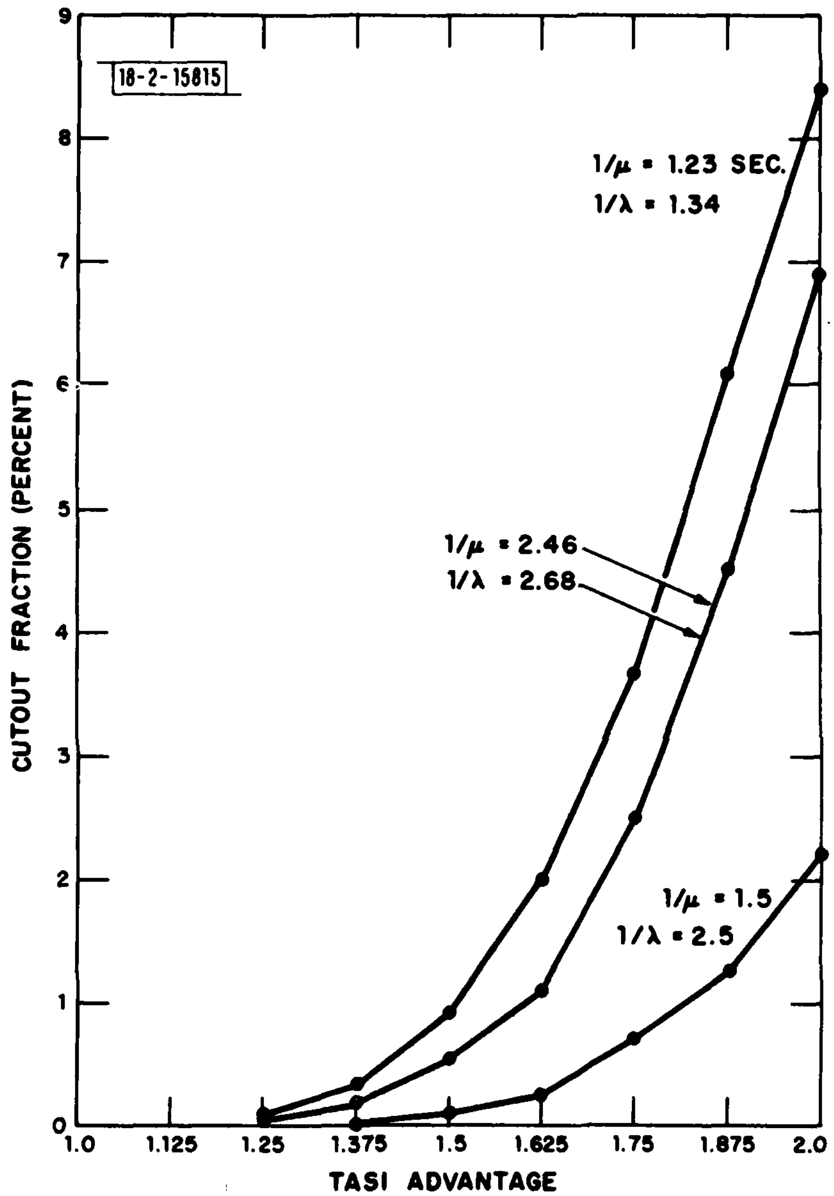

Fig.12. Effects of changes in mean talkspurt duration $1/\mu$ and mean silence duration $1/\lambda$. Other system parameters are the same as in Fig.9. Dynamic channel allocation, but no buffering, was used.

The sensitivity to the inter-packet interval was tested by doubling T_s from 20 to 40 msec while maintaining other system variables the same. Changes in performance results were negligible.

The number of speakers per node was varied from 5 to 20, and results are summarized in Fig. 13. As expected, increasing the number of callers per node produces more accurate prediction and better performance. However, dynamic allocation based on prediction produces improvements in all cases over fixed allocation.

VII. SUMMARY OF POTENTIAL TASI PERFORMANCE IMPROVEMENTS

A summary of potential TASI advantage improvements as determined by the simulations is presented in Fig. 14. Here system TASI advantage is plotted as a function of the number of off-hook callers at each node for various combinations of prediction and buffering. The results were obtained by requiring the cutout fraction not to exceed 0.5% and to determine at what TASI advantage this level of performance would be achieved in each case. For example, the results for 10 speakers/node are obtained from Fig. 11 by determining at what TASI advantages the various curves cross a cutout fraction threshold of 0.5%. As mentioned earlier, this is a conservative threshold for cutout fraction. As in Fig. 13, the satellite capacity is taken as 8M slots where M is the number of speakers per node. The assumed reservation response time was 280 msec as in most previously-presented cases.

Referring to the no buffering, fixed allocation case as a baseline, the various levels of performance improvement are apparent. Even for the

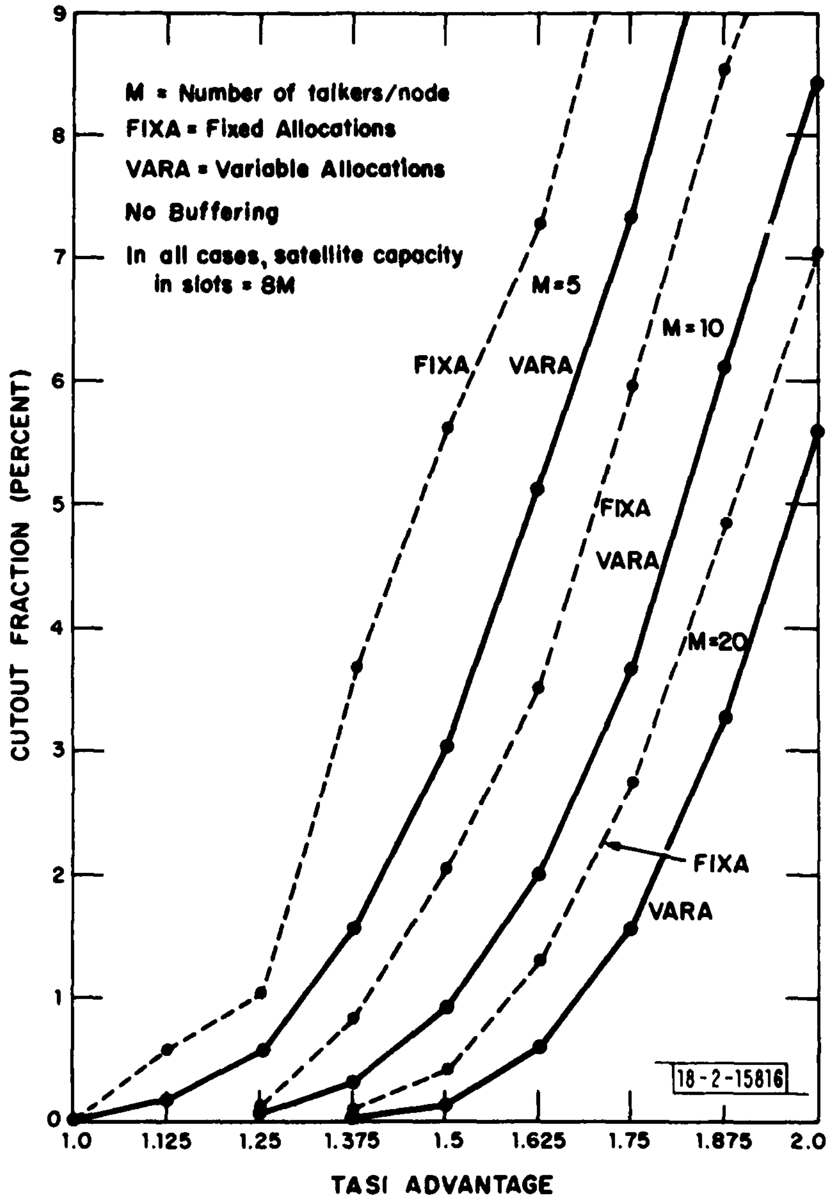

Fig.13. Effect of variation in the number of callers per node.

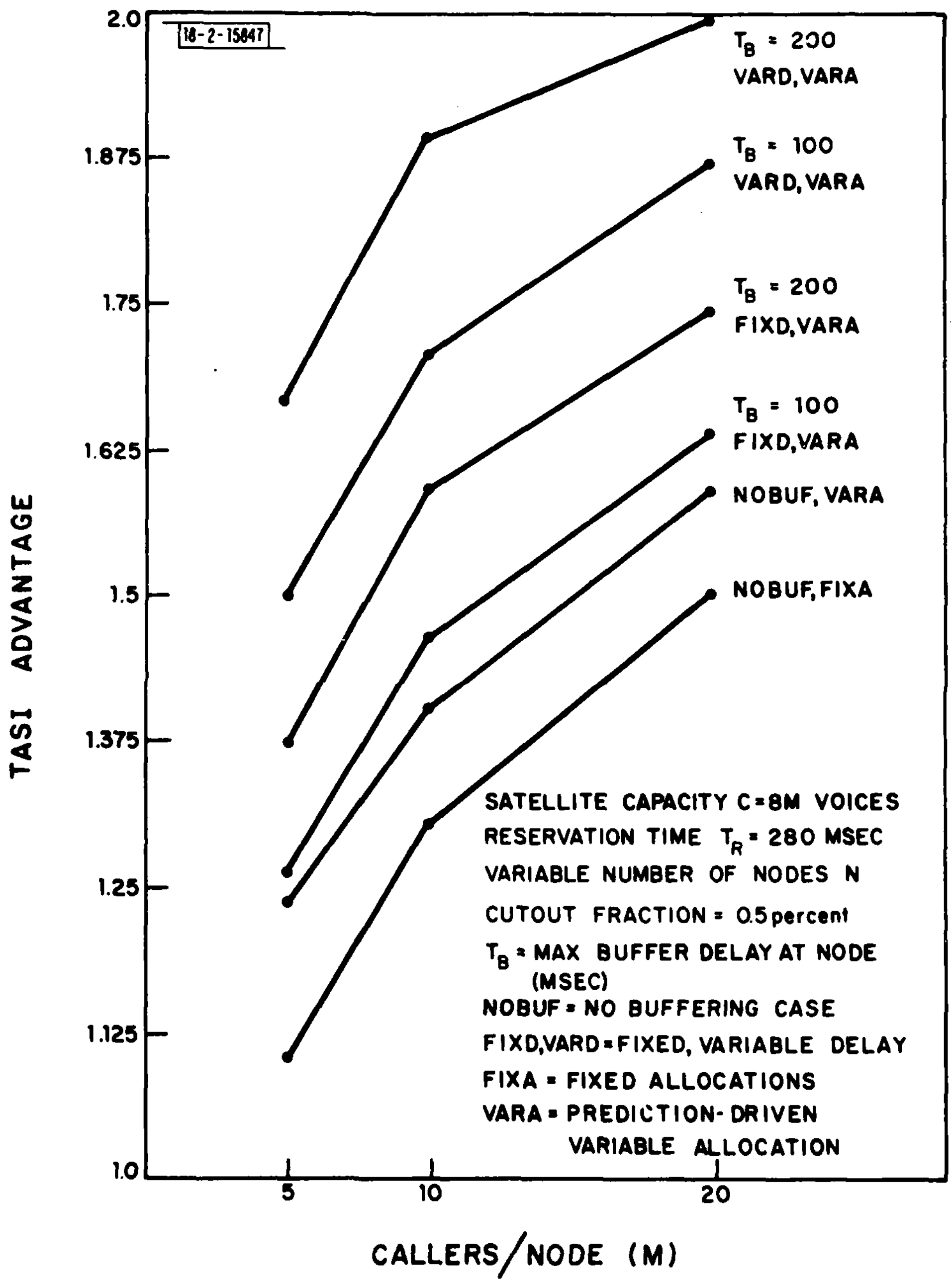

Fig.14. TASI advantage as a function of number of off-hook callers per node for various combinations of buffering and allocation strategies.

case of only 5 speakers/node, respectable values of TASI advantage can be achieved. The ordering of performance for various combinations of prediction and buffering follows the previous discussion regarding Fig. 11.

VIII. CONCLUSIONS

Prediction and buffering of digital speech streams has been shown to provide potential performance improvement in the statistical multiplexing of speech on a demand-assigned satellite channel, in the case where only a small number of users are multiplexed at each node. One can take advantage of this improvement either by accommodating more callers at a given cutout fraction or by providing a lower cutout fraction to a fixed number of users. Taking maximum advantage of prediction requires a rapidly responsive demand assignment algorithm capable of changing channel allocations within slightly more than a satellite round-trip time.

The simulations have shown that reservations for channel capacity should be based on prediction plus a correctly selected margin. The "optimal" margin was empirically determined to be the quantity which results in a system-wide reservation level that is approximately equal to the channel capacity. It is possible for the system to adaptively establish such a margin in a dynamic fashion by observing the system-wide reservation rate and making suitable adjustments to the margin currently being used by a node.

Additional performance improvement can be achieved by buffering packets before their transmission. This improves prediction by reducing

the required predict-ahead interval. In addition, variable-length buffering provides a smoothing action between temporary overloads and more quiescent time periods. Variable-delay buffering was shown to be more productive than fixed-delay buffering.

ACKNOWLEDGMENTS

The authors wish to acknowledge many comments and suggestions from T. Bially, J. W. Forgie, and R. Berger.

REFERENCES

1. R. D. Rosner, "Optimization of the Number of Ground Stations in a Domestic Satellite," EASCON 1975 Proceedings, pp. 64A.

2. J. D. Barnla and F. R. Zitzmann, "The SBS Digital Communications Satellite System," EASCON 1977 Proceedings.

3. I. M. Jacobs, R. Binder, and E. V. Hoversten, "General Purpose Packet Satellite Networks," Proc. IEEE, 66, 1448 (1978).

4. K. Bullington and J. M. Fraser, "Engineering Aspects of TASI," Bell System Technical Journal, 38, 353 (1959).

5. S. J. Campanella, "Digital Speech Interpolation," COMSAT Technical Review, 6, 127, (1976).

6. P. J. Brady, "A Technique for Investigating On-Off Patterns of Speech," Bell System Technical Journal, 44, 1 (1965).

7. P. J. Brady, "Effects of Transmission Delay on Conversational Behavior on Echo-Free Telephone Circuits," Bell System Technical Journal, 50, 115 (1971).

8. E. M. Hofstetter, "Speaker Activity Prediction," in Network Speech Processing Program Annual Report, Lincoln Laboratory, M.I.T. (30 September 1978) pp. 20-23, DDC AD-068318.

9. C. J. Weinstein and E. M. Hofstetter, "The Tradeoff Between Delay and TASI Advantage in a Packetized Speech Multiplexer," IEEE Trans. on Commun., COM-27, 1716 (1979).

10. R. Binder, "A Dynamic Packet Switched System for Satellite Broadcast Channels," Proc. ICC 1975, San Francisco, CA (June 1975).

11. J. W. Forgie, "Speech Transmission in Packet-Switched Store-and-Forward Networks," National Computer Conference Proceedings (1975).

12. I. Gitman and H. Frank, "Economic Analysis of Integrated Voice and Data Networks: A Case Study," Proc. of the IEEE, 66, 1549 (1978).

13. G. J. Coviello, "Comparative Discussion of Circuit- vs. Packet-Switched Voice," IEEE Trans. on Commun., COM-27, 1153 (1979).

14. J. W. Forgie and A. G. Nemeth, "An Efficient Packetized Voice/Data Network Using Statistical Flow Control," Proc. ICC 1977 (June 1977).

15. J. M. Elder and J. F. O'Neill, "A Speech Interpolation System for Private Networks," NTC Conference Record, pp. 14.6.1-14.6.5, Brimingham, Alabama (December 1978).

16. J. W. Forgie, "Network Speech System Implications of Packetized Speech," Lincoln Laboratory Annual Report to Defense Communications Agency (September 1976) DDC AD-A045455/3.

17. C. J. Weinstein, "Fractional Speech Loss and Talker Activity Model for TASI and for Packet-Switched Speech," IEEE Trans. on Commun. COM-26, 1253 (1978).

18. R. Syski, Introduction to Congestion Theory in Telephone Systems, (Oliver and Boyd, London, 1960, pp. 243-245.

19. L. Kleinrock, Queueing Systems, Vol. I: Theory, (Wiley, New York, 1975).

20. R. Berger, private communication.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: ESD-TR-79-249	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Prediction and Buffering of Digital Speech Streams for Improved TASI Performance on a Demand-Assigned Satellite Channel	5. TYPE OF REPORT & PERIOD COVERED: Technical Note
	6. PERFORMING ORG. REPORT NUMBER: Technical Note 1979-75
7. AUTHOR(s): William Kantrowitz, Clifford J. Weinstein, Edward M. Hofstetter	8. CONTRACT OR GRANT NUMBER(s): F19628-80-C-0002
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Lincoln Laboratory, M.I.T., P.O. Box 73, Lexington, MA 02173	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Program Element No. 33126K
11. CONTROLLING OFFICE NAME AND ADDRESS: Defense Communications Agency, 8th Street & So. Courthouse Road, Arlington, VA 22204	12. REPORT DATE: 15 November 1979
	13. NUMBER OF PAGES: 48
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Electronic Systems Division, Hanscom AFB, Bedford, MA 01731	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

TN-1979-75

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Demand Assignment Multiple Access (DAMA)
Time-Assigned Speech Interpolation (TASI)
speaker activity prediction
stream reservations
packet speech
Speech Activity Detection (SAD)
Digital Speech Interpolation (DSI)
cutout fraction
speech buffering

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An approach for achieving efficient TASI-like multiplexing of speech on a demand-assigned satellite channel, where a number of ground stations each support a small number of off-hook callers, is proposed and evaluated. The approach presupposes a demand-assignment scheme which allows rapid changes in channel capacity assignments, and includes channel reservation requests based on speaker activity prediction, and buffering of speech at the nodes to increase multiplexing efficiency. System performance results obtained by computer simulation indicate potentially significant reductions in speech cutout fraction and accompanying increases in system TASI advantage.

DD FORM 1473, 1 JAN 73 — EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

207650 GM